Supporting Information:

Pairwise interactions between linear alkanes in water measured by AFM force spectroscopy

*Chad Ray, Jason R. Brown, Andrea Kirkpatrick[†], and Boris B. Akhremitchev**

Department of Chemistry, Duke University, Durham, North Carolina 27708

* e-mail: boris.a@duke.edu

This Supporting Information provides a detailed description of the data manipulation and includes all of the histograms and fits referred to in the main text.

Data Collection and the Extended Freely Jointed Chain Fit:

Upon collection, force curves are saved as waves in Igor Pro 5 (Wavemetrics Inc., Portland, OR). These waves are processed with a custom program written in Matlab (MathWorks, Inc., Natick, MA). The majority of the collected force curves do not exhibit characteristic separation events. Data processing is automated to remove user bias and speed up the data analysis. The program converts the force curves from deflection versus displacement into force versus separation. The deflection noise value is calculated as a standard deviation from the off-surface part of the approach line, and the mean noise value based threshold is used to detect rupture events. A threshold of five standard deviations is applied to the retract curve to detect the abrupt

† Current address: Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125

transitions in force. After this simple analysis, a filter is applied to remove the rupture transitions that occur at separations too close to the surface (<10 nm). A second set of filters refines parameters prior to the freely jointed chain fit to select force events that have a force pattern typical for separation events coupled to a stretching polymer (when gradual tensioning of the polymeric tether is followed by the abrupt release of the accumulated stress). Our tests indicate that this initial processing does not eliminate force curves that can be considered for further analysis by a trained user. The selected polymer-stretching events are fit with the extended freely jointed chain model[1] to extract the contour length and Kuhn length parameters and determine the loading rate for each separation event. This model is an extension of the commonly used freely jointed chain (FJC) model. The FJC model predicts extension of the polymer chain $x(F)$ with the Kuhn length l_k and contour length L_c as a function of applied force F according to

$$x(F) = L_c \cdot \left[\coth\left(F \cdot l_k \Big/ k_B \cdot T \right) - k_B \cdot T \Big/ F \cdot l_k \right] \qquad \text{(S1)}$$

where k_B is the Boltzman's constant and T is the temperature. This equation is usually written as $x(\beta) = L_c \cdot L(\beta)$ were $L(x)$ is the Langevin function and $\beta = F \cdot l_k / (k_B \cdot T)$.

Besides an entropic elasticity of the polymer chain included in FJC model, the extended model includes elongation of the PEG chain due to monomer elasticity as well as conformational transition between helical and planar conformations of the PEG chain in aqueous solutions.[1] In this model, the contour length of stretched polymer consists of the lengths of polymer segments at two different conformations, planar and helical:

$$L_c = N_{planar} \cdot L_{planar} + N_{helical} \cdot L_{helical} \qquad \text{(S2)}$$

Here N_{planar} and $N_{helical}$ are the numbers of segments in planar and helical conformations respectively. L_{planar} and $L_{helical}$ are the corresponding monomer lengths that are fixed to 3.58 Å and 2.8 Å respectively in our calculations.[1] Contour length defined by equation S2 can be related to the common definition of contour length (the maximum distance between ends of the linear polymer chain) by noting that if $N_{helical}$ and N_{planar} have fixed (force-independent) values then the usual definition of contour length can be applied. The ratio of $N_{helical}$ to N_{planar} depends on applied force according to:

$$\frac{N_{helical}}{N_{planar}} = e^{\Delta G(F)/k_B \cdot T}$$
$$\Delta G(F) = \Delta G_0 - F \cdot \left(L_{planar} - L_{helical} \right) \qquad \text{(S3)}$$

Here $\Delta G(F)$ is the force-dependent free energy difference between the two states and ΔG_0 is this difference at zero applied load, fixed to 7.48 kJ/mol in our calculations.[1] The overall PEG chain with N monomers the extension is

$$x(F) = N \cdot \left(\frac{L_{planar}}{e^{+\Delta G(F)/k_B \cdot T} + 1} + \frac{L_{helical}}{e^{-\Delta G(F)/k_B \cdot T} + 1} \right) \cdot \left[\coth(\beta) - 1/\beta \right] + N \cdot \frac{F}{K_S} \qquad \text{(S4)}$$

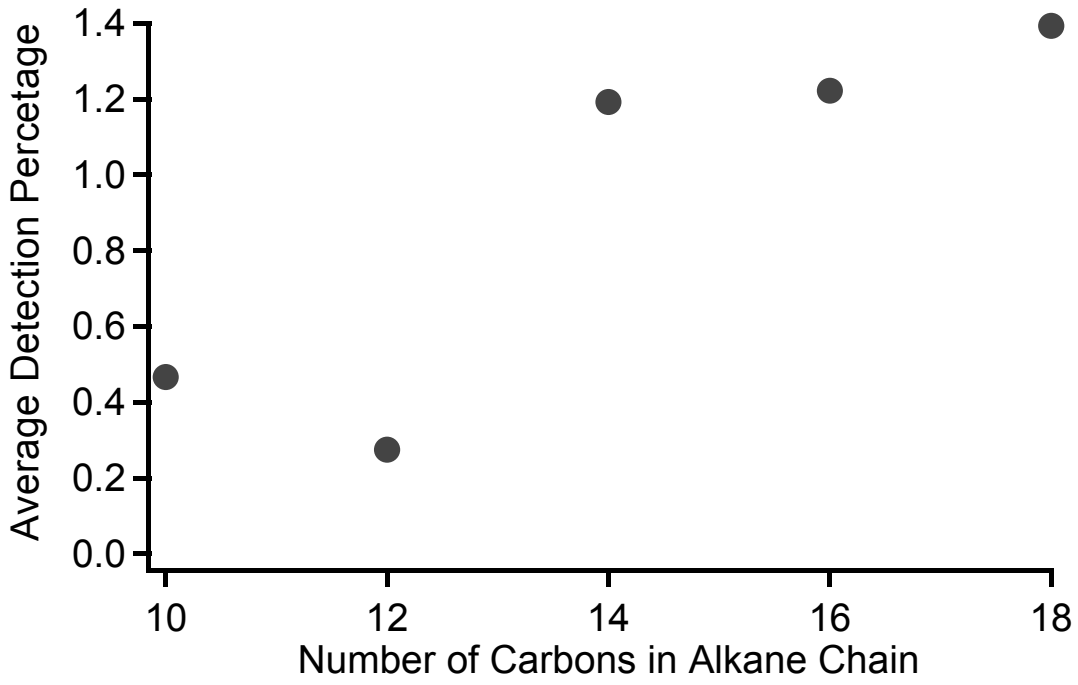
Here the segmental elasticity K_s provides the chain extension at high loads and is held at 150 N/m[1] and the other parameters are described above. This model was used to fit the force curves with two free parameters: the number of monomers N in the chain and the Kuhn length l_k. The Kuhn length was allowed to vary to obtain a close fit to the data near the separation point. The tether spring constant was obtained as the slope of the fit curve at the rupture point. It was noted that when both FJC and the extended FJC models were used to fit the experimental data, the FJC model produced systematically higher tether elasticity values. The systematic error in tether elasticity will propagate in the error in the loading rate that is calculated according to

$$l_r = v\,(k_t^{-1} + k_c^{-1})^{-1} \qquad \text{(S5)}$$

where k_c is the spring constant of the cantilever and v is the velocity of the cantilever base. The extended FJC model fits the stretching curve closely, providing more accurate loading rate determination.

Detection probability of rupture events as a function of the alkane chain length

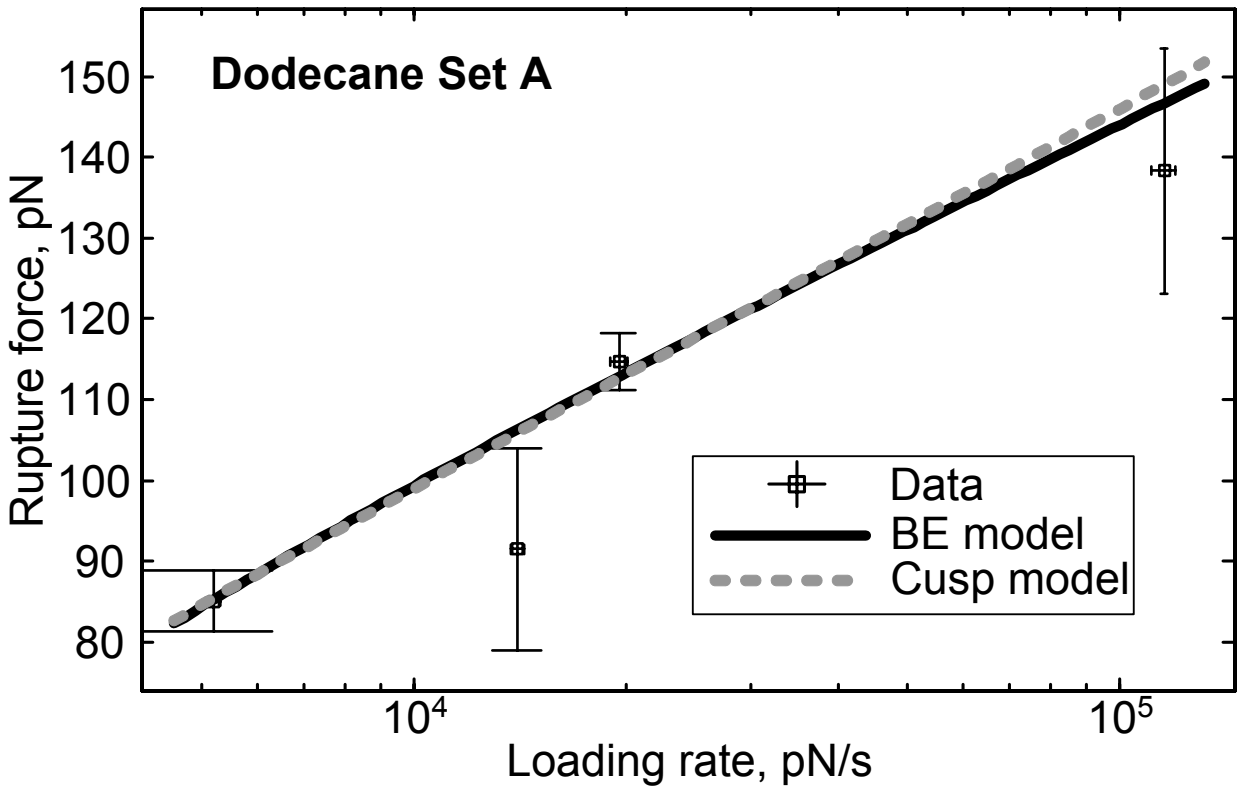
Most contacts of the AFM tip with the surface do not yield a rupture event. The detection probability is the ratio of selected double-tether force events to total surface approaches. The detection probability averaged for different probe velocities is shown in the figure below as a function of the alkane chain length. It can be noticed that decane and dodecane yielded much fewer force events than the larger alkane chains.

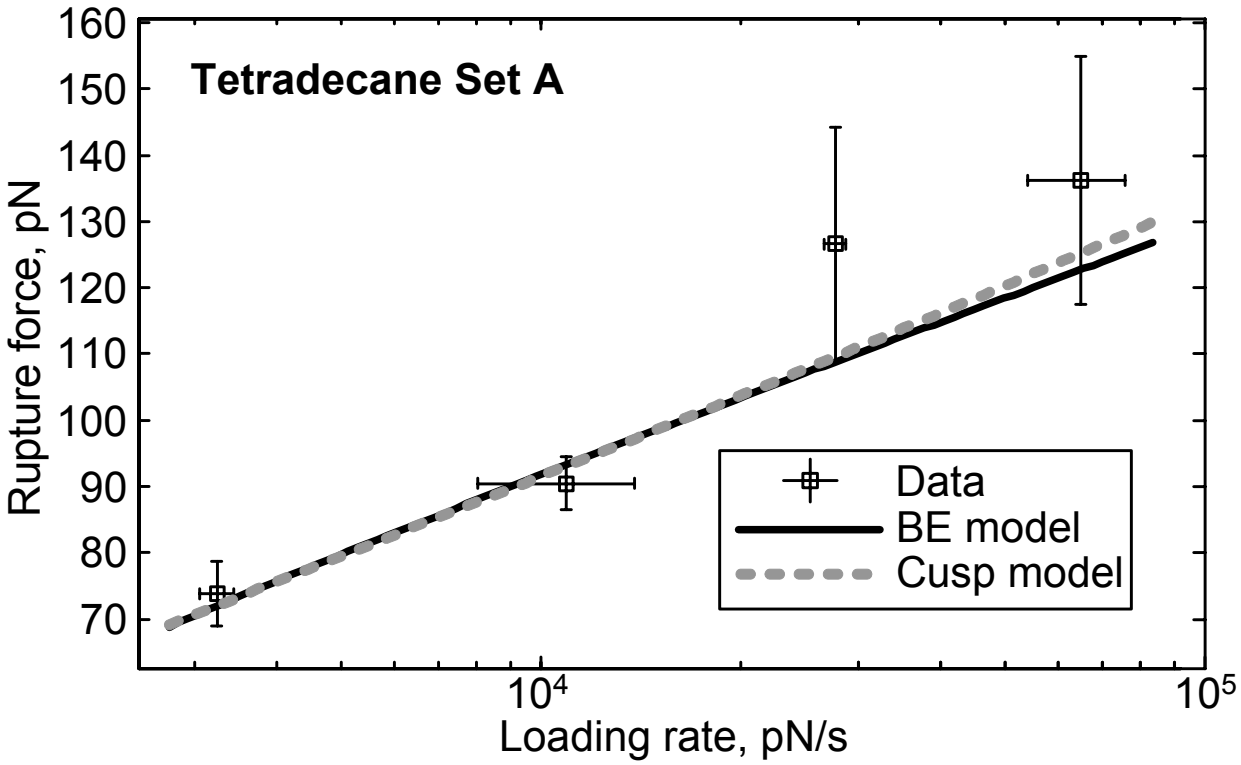


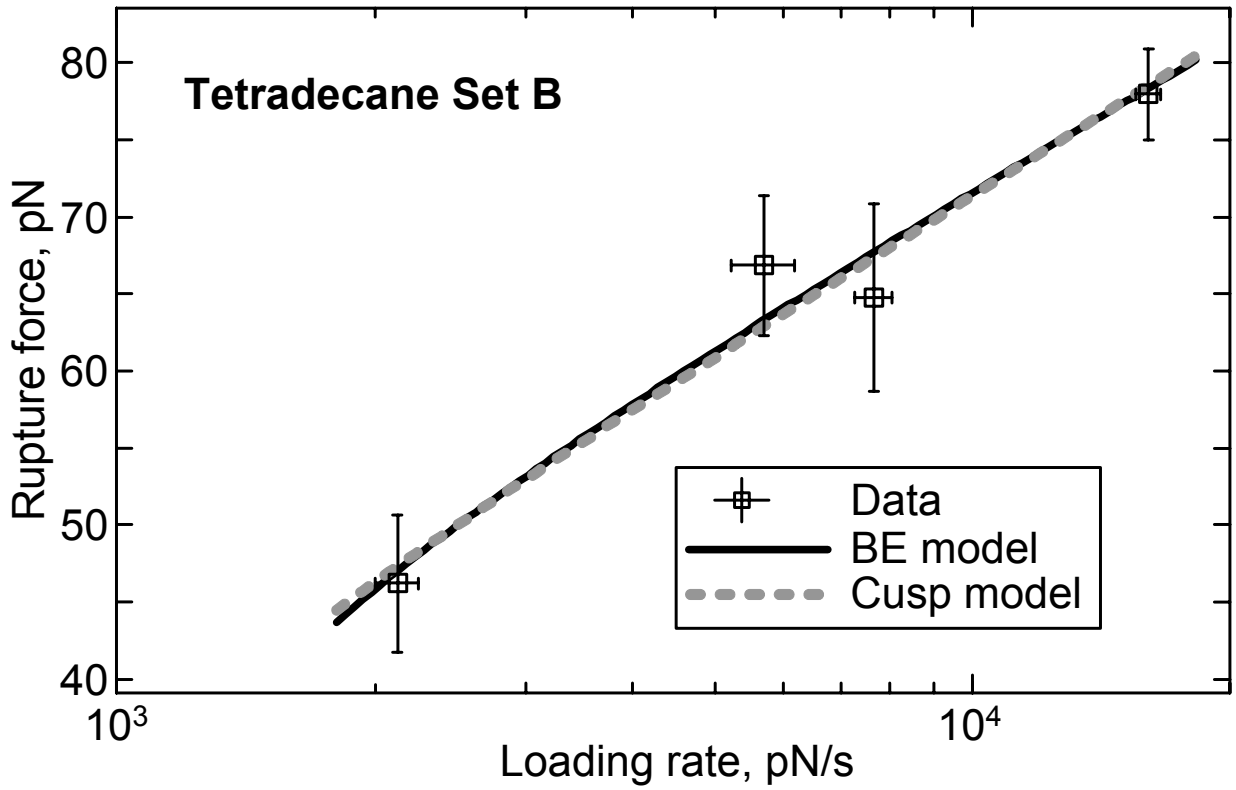
Data analysis using the Most Probable Force and the Most Probable Loading Rate:

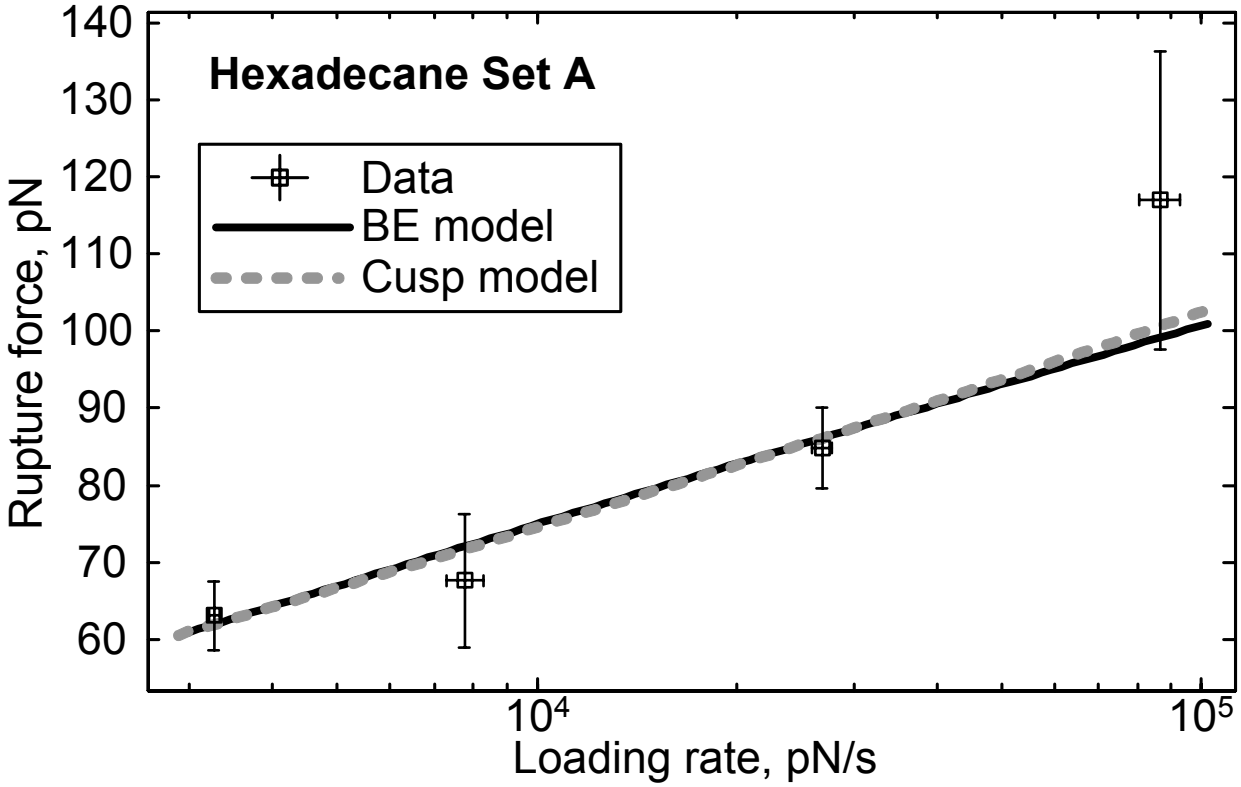
In finding the most probable force from the histograms of experimental rupture forces two important effects have to be considered. The first is limited force sensitivity and the second is the presence of the high forces tail in the distribution.[2] The limited force sensitivity is accounted for by applying a window function.[2] Because the exact shape of the window function is unknown, a Gaussian error function is selected for convenience. The position and width of the window function are selected to match the rising edge of the force distribution[3] and kept the same for all fitted histograms. Each force distribution is fit with a Gaussian curve multiplied by the window function to determine the most probable force. This is done in lieu of fitting with the individual bond probability distribution function because the Bell-Evans model (this model is commonly used to calculate the distribution of rupture forces) predicts a sharp cutoff of forces at the high force limit while the experimental data contain a high force tail, as described in the main text. If fit by the Bell-Evans probability distribution function, the resulting most probable force would be shifted improperly to higher forces as a result of the high force tail found in experimental data (as seen in the figures below). To sidestep this problem, a Gaussian curve is fit to the experimental force distribution because the Gaussian function can fit the high forces tail. The fitting minimized root-mean-square error between the fit function and the histogram by adjusting three free parameters: the center of the Gaussian curve, its width and the amplitude. The most probable force is taken as the center of the Gaussian curve.

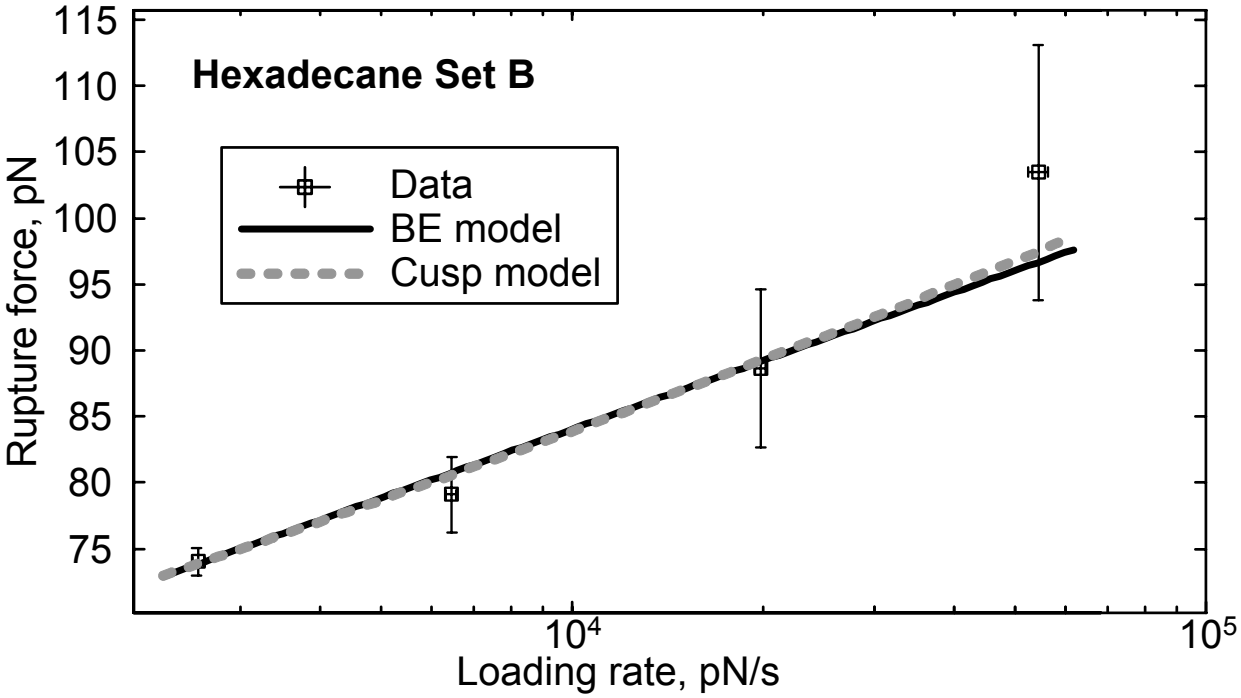
The most probable force vs. loading rate dependences were fit using kinetic models as described in the main text. Fits to the Bell-Evans (solid black line) and cusp (dotted grey line) models are shown below for each of the alkanes. The random errors in the most probable forces are determined from the covariance matrix.[4]

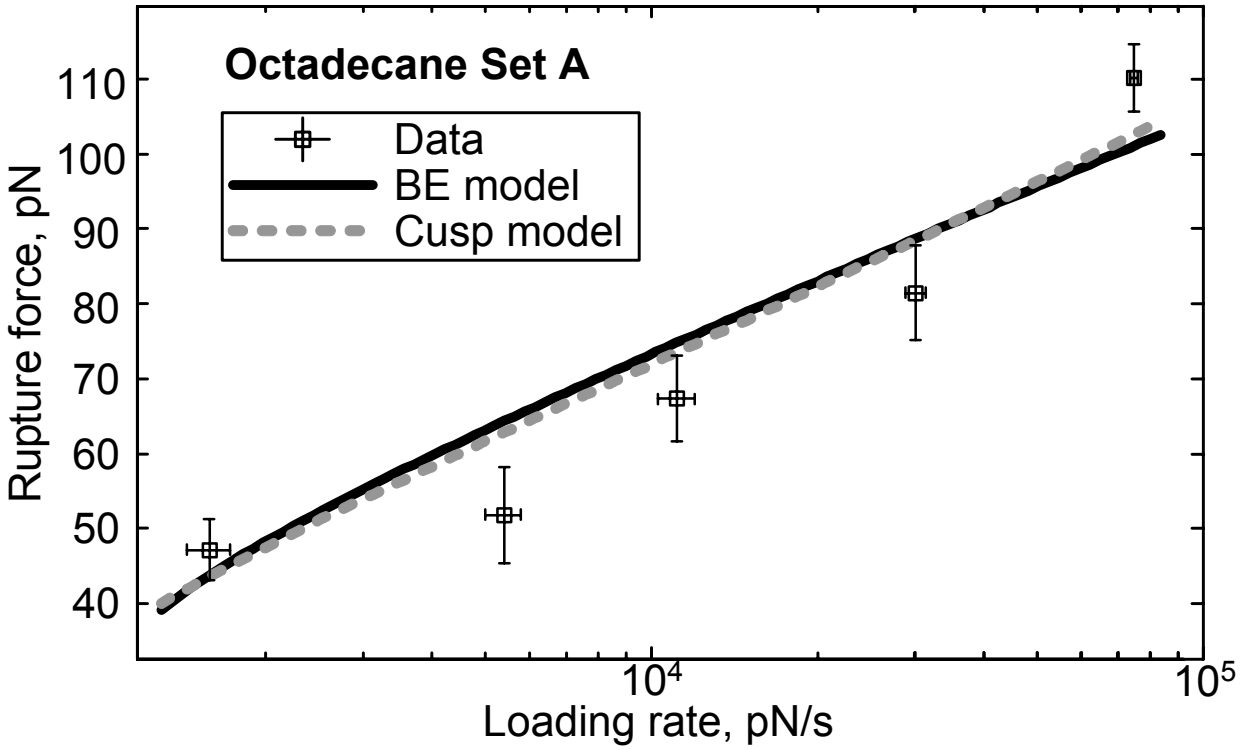

Dodecane Set A
Rupture force, pN
150
140
130
120
110
100
90
80
10^4
10^5
Loading rate, pN/s
Data
BE model
Cusp model

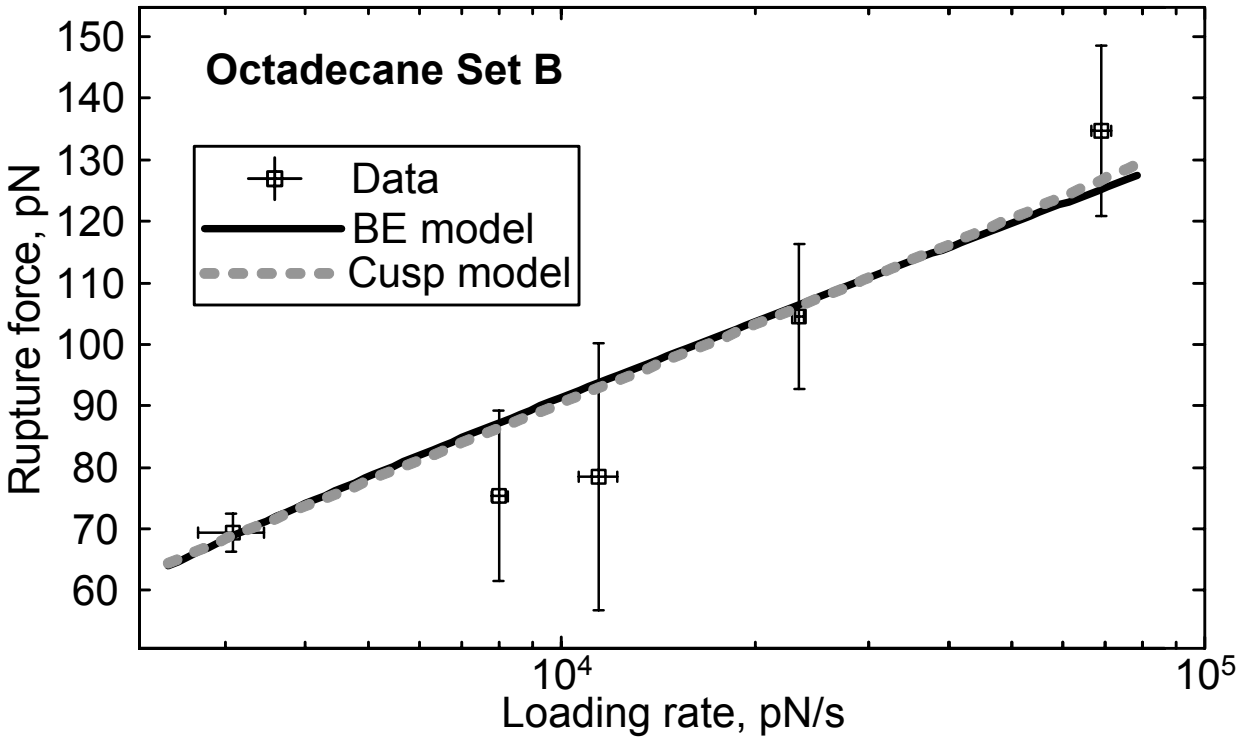

Tetradecane Set A
Rupture force, pN
Loading rate, pN/s
Data
BE model
Cusp model


Tetradecane Set B
Rupture force, pN
Loading rate, pN/s
Data
BE model
Cusp model


Hexadecane Set A
Data
BE model
Cusp model
Rupture force, pN
Loading rate, pN/s
140
130
120
110
100
90
80
70
60
10^4
10^5

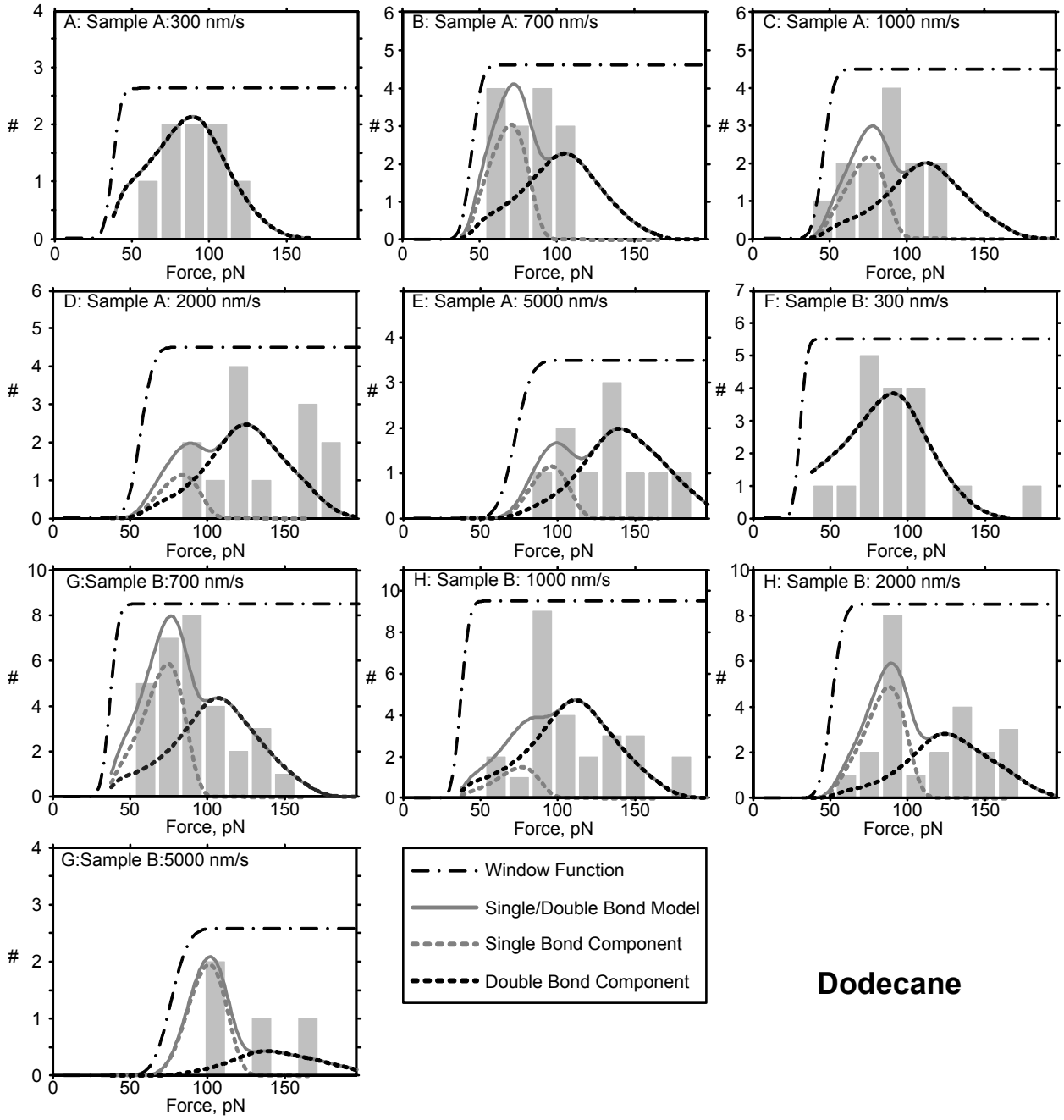

Hexadecane Set B
Data
BE model
Cusp model
Rupture force, pN
Loading rate, pN/s
115
110
105
100
95
90
85
80
75
10^4
10^5

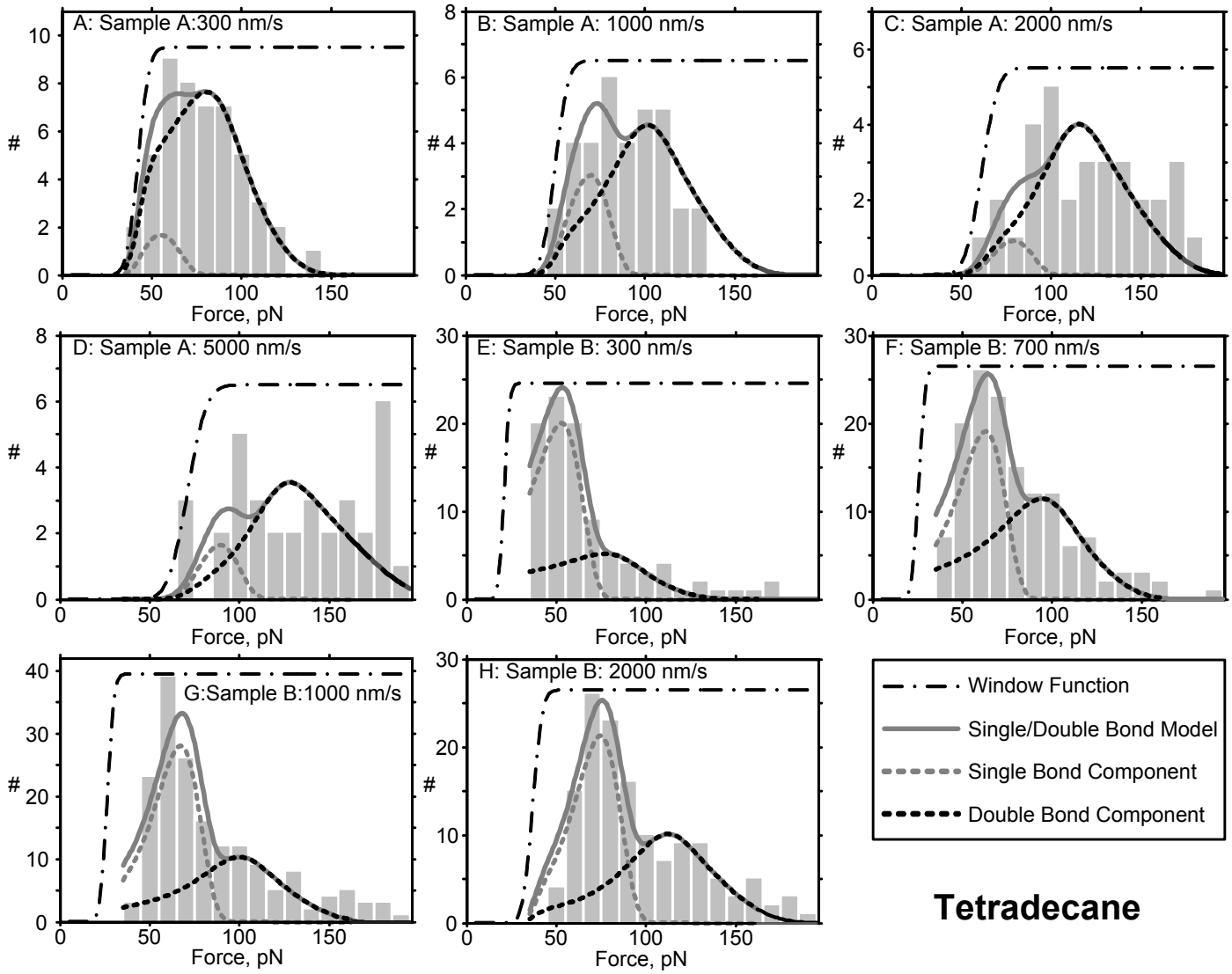

Octadecane Set A
Data
BE model
Cusp model
Rupture force, pN
110
100
90
80
70
60
50
40
10^4
10^5
Loading rate, pN/s

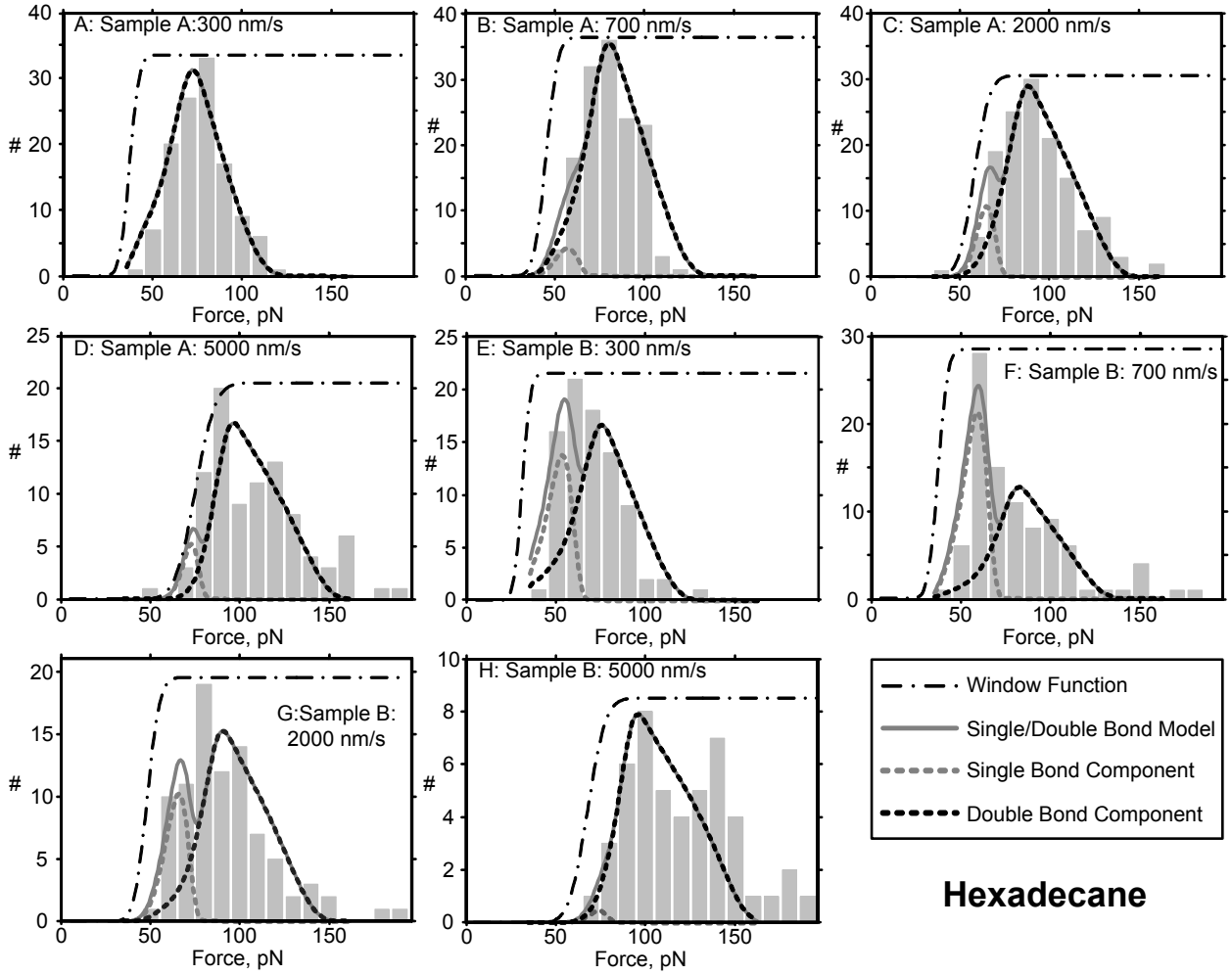

Octadecane Set B
Data
BE model
Cusp model
Rupture force, pN
150
140
130
120
110
100
90
80
70
60
10^4
10^5
Loading rate, pN/s

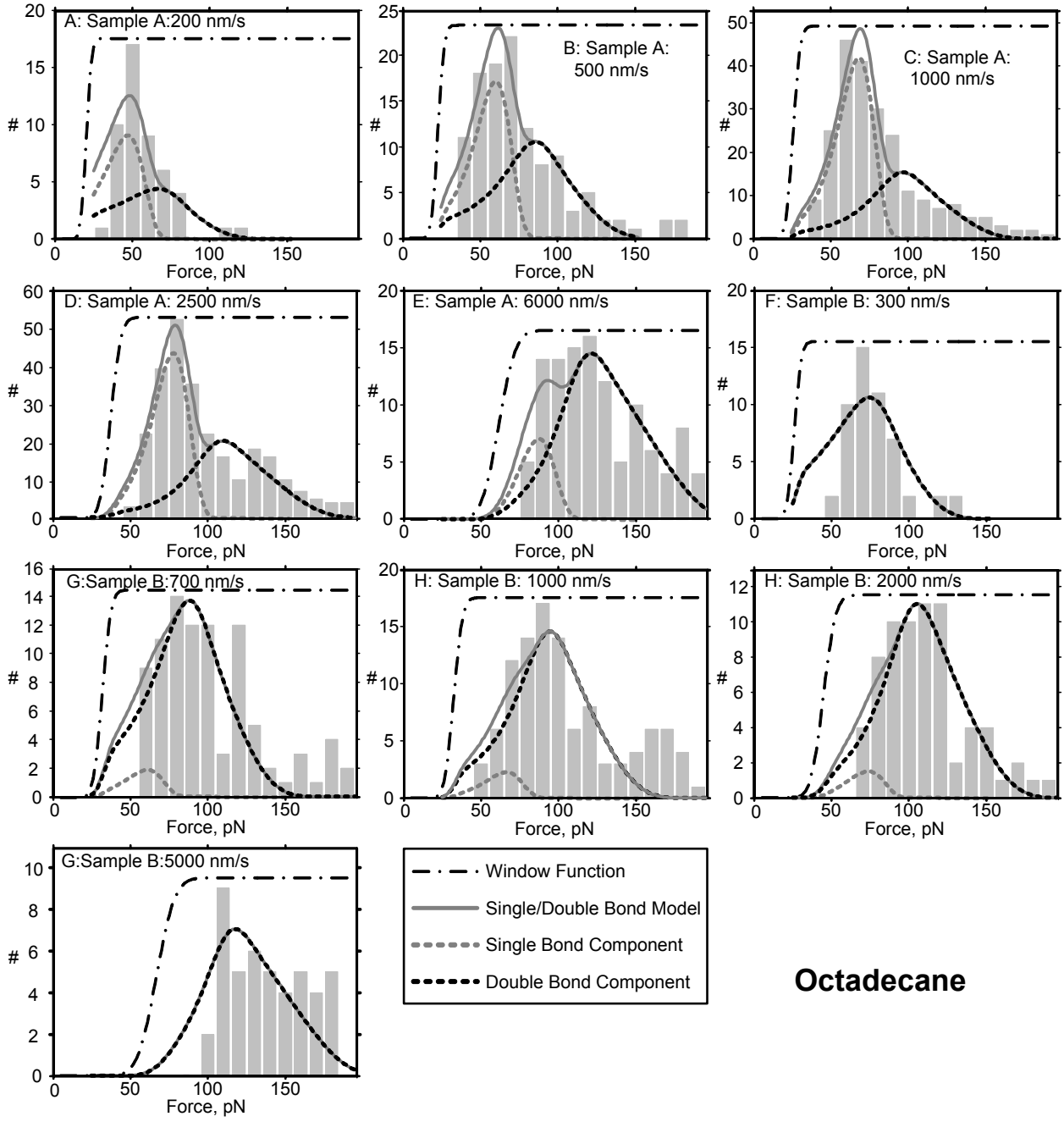
Determination of the Most Probable Force and the Most Probable Loading Rate:

Below the experimental rupture force histograms for each set of the alkanes are shown with the cumulative fit (solid grey line), as well as its components (individual bond component, dotted grey line and two-bond component, dotted black line). The bin size for all force histograms is held the same for all histograms and is equal to 15 pN. The black dash-dotted line is the window function fit, as mentioned above, to account for limited force sensitivity. This window function is scaled by the height of the histogram for clarity. Kinetic parameters from these fits are given in the main text.


A: Sample A:300 nm/s
B: Sample A: 700 nm/s
C: Sample A: 1000 nm/s
D: Sample A: 2000 nm/s
E: Sample A: 5000 nm/s
F: Sample B: 300 nm/s
G:Sample B:700 nm/s
H: Sample B: 1000 nm/s
H: Sample B: 2000 nm/s
G:Sample B:5000 nm/s
#
Force, pN
Window Function
Single/Double Bond Model
Single Bond Component
Double Bond Component
Dodecane


A: Sample A:300 nm/s
B: Sample A: 1000 nm/s
C: Sample A: 2000 nm/s
D: Sample A: 5000 nm/s
E: Sample B: 300 nm/s
F: Sample B: 700 nm/s
G:Sample B:1000 nm/s
H: Sample B: 2000 nm/s
#
Force, pN
Window Function
Single/Double Bond Model
Single Bond Component
Double Bond Component
Tetradecane


A: Sample A:300 nm/s
B: Sample A: 700 nm/s
C: Sample A: 2000 nm/s
D: Sample A: 5000 nm/s
E: Sample B: 300 nm/s
F: Sample B: 700 nm/s
G:Sample B: 2000 nm/s
H: Sample B: 5000 nm/s
Force, pN
#
Window Function
Single/Double Bond Model
Single Bond Component
Double Bond Component
Hexadecane


A: Sample A:200 nm/s
B: Sample A:
500 nm/s
C: Sample A:
1000 nm/s
D: Sample A: 2500 nm/s
E: Sample A: 6000 nm/s
F: Sample B: 300 nm/s
G:Sample B:700 nm/s
H: Sample B: 1000 nm/s
H: Sample B: 2000 nm/s
G:Sample B:5000 nm/s
#
Force, pN
Window Function
Single/Double Bond Model
Single Bond Component
Double Bond Component
Octadecane

References:

(1) Oesterhelt, F.; Rief, M.; Gaub, H. E. *New Journal of Physics* **1999**, *1*, 6.1.

(2) Friedsam, C.; Wehle, A. K.; Kuhner, F.; Gaub, H. E. *Journal of Physics-Condensed Matter* **2003**, *15*, S1709.

(3) Ray, C.; Akhremitchev, B. B. *Journal of the American Chemical Society* **2005**, *127*, 14739.

(4) Bevington, P. R.; Robinson, D. K. *Data Reduction and Error Analysis for the Physical Sciences*, 2 ed.; McGraw-Hill: New York, 1991.